DOCUMENT RESUME

ED 022 454 JC 680 328

By-Kilpatrick, Gordon

A CONSIDERATION OF PLACEMENT TESTING.

El Camino Coll., El Camino College, Calif.
Report No-OIR-ECC-68-2
Pub Date May 68
Note-15p.
EDRS Price MF-$0.25 HC-$0.68
Descriptors-ACADEMIC ABILITY, COLLEGE STUDENTS, COMPUTERS, DATA PROCESSING, *EDUCATIONAL TESTING, ENGLISH PROGRAMS, *JUNIOR COLLEGES, *PREDICTIVE ABILITY (TESTING), PRETESTING, STUDENT ABILITY, *STUDENT PLACEMENT, STUDENT TESTING, *TESTING PROGRAMS
Identifiers-Cooperative English Test, Purdue Placement Test in English, *School and College Ability Tests

This report consists of three sections, the first of which is a discussion of placement testing. It is maintained that the strengths and weaknesses of a tested population cannot be determined by a single composite test--yet, many institutions place students in beginning courses on the basis of such limited information. To assess basic capabilities assumed to be acquired in the lower level course, a test must be devised after the course objectives are known. National standardized tests are not always sufficient. By employing multiple correlations or multiple regressions (which are explained in more detail in section two of the report), the proper placement of a student can be enhanced. Data can be computer processed for the selection of the best predictors in rank order. It is suggested that regression equations developed from data collected in one testing session be applied to the data collected during the testing one calendar year later. Problems such as the contamination of input data (e.g. different grading standards used by various teachers) are discussed with suggestions being given for developing more accurate regression equations. Section three describes English placement at El Camino College (Calif.). Test scores from the Purdue Test, the Cooperative English Test, and SCAT were submitted to a computer which ranked the test scores in order of their effectiveness as predictors. Recommendations for the El Camino program are given. (DG)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

OIR-ECC 68-2

A CONSIDERATION OF PLACEMENT TESTING

By

Gordon Kilpatrick

El Camino College

El Camino College, California

May, 1968

UNIVERSITY OF CALIF.
LOS ANGELES

JUN 18 1968

CLEARINGHOUSE FOR
JUNIOR COLLEGE
INFORMATION

ED022454

JC 680348

DISCUSSION

Presumably, placement tests are used to assess students' capabilities and subsequently place them into the highest level course commensurate with demonstrated competence in the skills and knowledge which are taught in the lower, prerequisite courses. To do less wastes the time and energies of both the student and teacher and the money of the school's sponsors. The decision made by the placement tests is a critical one, since an irreplaceable semester or quarter of an individual's productive life is involved. The loss of underplacement is obvious. Overplacement also involves a potential loss if the higher course presumes upon a developed background which is non-existent; current doctrine makes the resultant academic failure an irrevocable and unforgivable act which follows one via permanent transcripts to his grave. Placement *via* testing is not a subject to be treated lightly.

It is virtually impossible to identify the strengths and weaknesses of a tested population via a single total score on a composite test, especially when (1) the relative merit of the instrument is unknown when compared to other available instruments as a predictor of academic performance, and (2) the relative validity of the selected cutoff score is also unknown. A student's strengths disguise his weaknesses, and it is possible for two or more students having vastly different weaknesses to receive the same total score. It is unfortunate that students currently are customarily placed in beginning college courses of English, mathematics, chemistry, and others via a single score on a single placement test. Regardless of the extent of validity of the particular test, a single score can give only a restricted viewpoint of one rather broad, undetailed aspect of an individual student's capability. As a result, instruc-

tional efforts are not necessarily brought to bear directly on the student's weaknesses, yet to be identified.

Placement tests do not always measure the capabilities we assume for proper placement. Research has supported this contention by revealing low linear correlations between single test scores and the performance (letter grade received) in the course in which placed. Logically, a placement test should be designed to test for the presence of those capabilities assumed to be acquired in the lower level, possibly prerequisite, course which may be by-passed as a result of the test score. This requires a definitive statement of the course objectives of the course being considered for exemption, with the test *subsequently* being constructed based on that information. The wholesale use of national standardized tests is not warranted.

Until such placement tests related to curriculum content are constructed, a method of placing students in their beginning course should be investigated which places the student using more than a single aspect of his capabilities. It has been found that correlation between placement scores and academic performance in the course can be improved by resorting to multiple correlation. If a multiple regression equation is developed stepwise, the best predictors are easily identified. In this method, any quantitative data which could logically have bearing on prediction of actual performance in the course is considered. Computational requirements make computer usage extremely attractive. The BMD02R (Stepwise Regression) program used in the Western Data Processing Center's IBM Model 360/75 computer is appropriate. In this program, if the data are presented in rudimentary form, broken down as far as possible into test section subscores, the computer will select the best predictors *in rank order* to the neglect of poorer predictors.

Inasmuch as the dependent variable (letter grade in the course) is not available until at least a semester after the test is given and testing populations vary according to time of year, the best approach might be to develop the regression equations from data collected in one testing session and apply the equations for placement to the data collected during the testing session one calendar year later. If intervening testing populations prove typical, the same equations may be used throughout the year after development on the testing subpopulation which preponderates. Consideration may be given to development of the equations upon a representative sample of testing sessions interspersed throughout the year, but the staggered availability of final grades may prevent timely application to the next oncoming testing population.

Collecting the data constituting the large number of independent variables to be considered for possible inclusion in the regression equation may cause a somewhat long and fatiguing testing session. It will probably be desirable to constitute a representative sample from the testing population in order to minimize the number of students subjected to the increased strain and duress. The sample should be large enough to provide approximately 100 students completing each course with letter grades and for which a regression equation is to be developed.

An obvious contamination of the input data is the teacher effect upon the dependent variable. This is probably the weakest point of the entire method. Teachers do not teach alike nor grade alike. Some tend to grade more extremely than others, either generously or harshly. Some teachers' grades cluster at various points on the A to F continuum. Others cover the entire continuum with a somewhat even spread or maybe a normal distribution. If a list of prior residuals is available, they may be averaged for each teacher

and the prediction may be somewhat improved by applying an arithmetic correction to the dependent variable before developing the regression equation.

Once the regression equations are developed stepwise, a decision will have to be made regarding how many variables can be used feasibly for future mass testing after considering the ramifications of testing, scoring, data availability, and data format. The equation giving the highest *feasible* correlation should be used. Conceivably, the future placement test may consist of selected sections of two or more independent tests, the scores on those sections having been proved the best predictors. The selected equation may also require non-test data such as high school grade point average (overall or in selected cogent courses). The capability of retrieving such data from students' records will dictate its feasibility. Student records on magnetic tape will facilitate the use of this variable.

Naturally, the dependent variable must be presented in numerical form for computational purposes. The common grade point average numerical equivalents are recommended whereby A=4, B=3, C=2, D=1, and F=0.

For placement the highest entry course must be tested first. The appropriate equation is applied and academic performance is predicted. If success is defined as a minimum grade of "C" or better (2.0 or above), any predicted grade of "C-" or below (1.5 or less) would result in placement in the lower level course. This cutoff score is subject to adjustment upward to offset regression resulting from limitations on range of talent resulting from previous placement. If three entry courses are available, the regression equation appropriate to the second level course would then be applied to the available data and a prediction of lack of success in that course would result in the student's placement in the third level entry course.

Appendix I

CORRELATION AND REGRESSION

Correlation exists between two sets of data if knowledge of one set allows reasonably accurate prediction of the corresponding values constituting the other set. If increase in one value causes a *proportional* increase in its corresponding value in the matched data set, the prediction can be made with 100% accuracy and an index of degree of correlation (correlation coefficient) takes the maximum value of (+)1. If an increase in one value causes a proportional *decrease* in its corresponding value, perfect prediction can still be made, but *negative* correlation exists with a correlation coefficient having the opposite extreme limiting value of -1.

A graphical plot of two sets of perfectly correlated scores reveals a straight line-like (*linear*) shape. If a line, known as a *regression line*, is drawn through the points known and the perfect correlation is assumed to hold, then knowledge of one set of values, the *independent* values (X_1, X_2, X_3, ... X_n) allows prediction of the corresponding *dependent* values (Y_1, Y_2, Y_3, ... Y_n). See Figure 1.

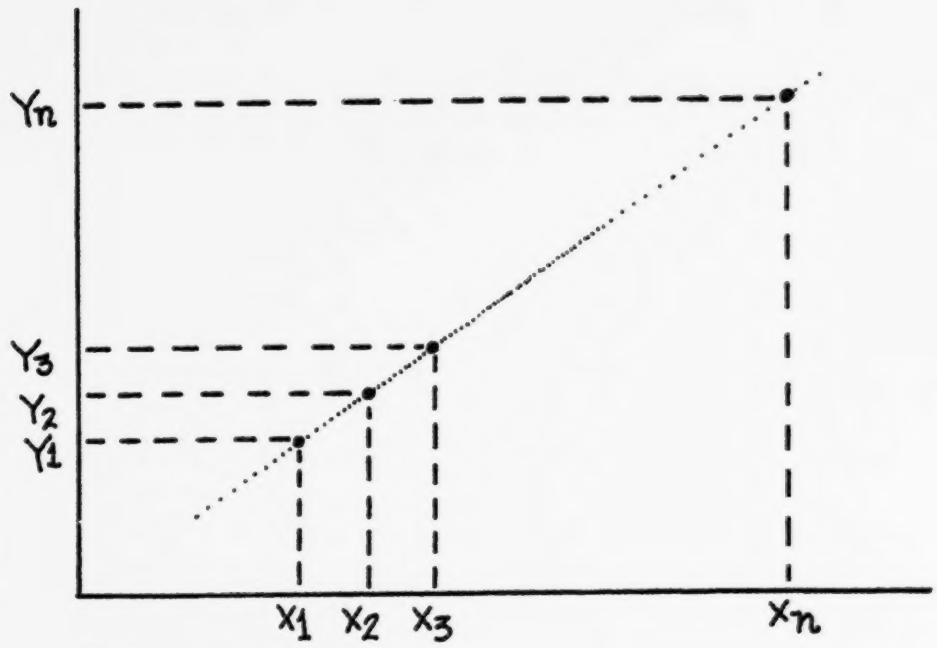

Figure 1

A mathematical equation known as a *linear regression equation* may be derived which describes the line and allows direct substitution of the known independent values in order to predict the corresponding matched (but unknown) dependent values.

Data sets are seldom perfectly correlated. Correlation coefficients vary between the two extreme limiting values of +1 and -1 which indicate perfect positive and negative correlation respectively. A correlation coefficient equalling zero means no correlation whatsoever. Regression lines may still be constructed with the errors contributed by deviation of the paired value points from the line being minimized. This may be done graphically by measuring the distance of the point from the line, squaring the distance, and summing the squares. The line which has the minimum sum of squares is said to be the "*least squares fit*." A mathematical procedure can provide the equation of this line without either the tedium or the trial-and-error approach of the graphical method. The magnitude of total error of prediction will exceed the accuracy of prediction unless the absolute value of the correlation coefficient equals or exceeds 0.71.

Regression occurs when predictions are made with non-perfect correlations. The result of this is that the predicted value approaches the mean of the dependent variables as the correlation coefficient approaches zero. With zero correlation, the regression line is horizontal which allows prediction of only the mean of the dependent variables used in establishing the level of correlation. See Figures 2 and 3.

Limitations on the number of cases being considered resulting from application of prior selection standards will contribute to a *limited range of talent* condition. The effect will be a lowered correlation coefficient and regression will occur.

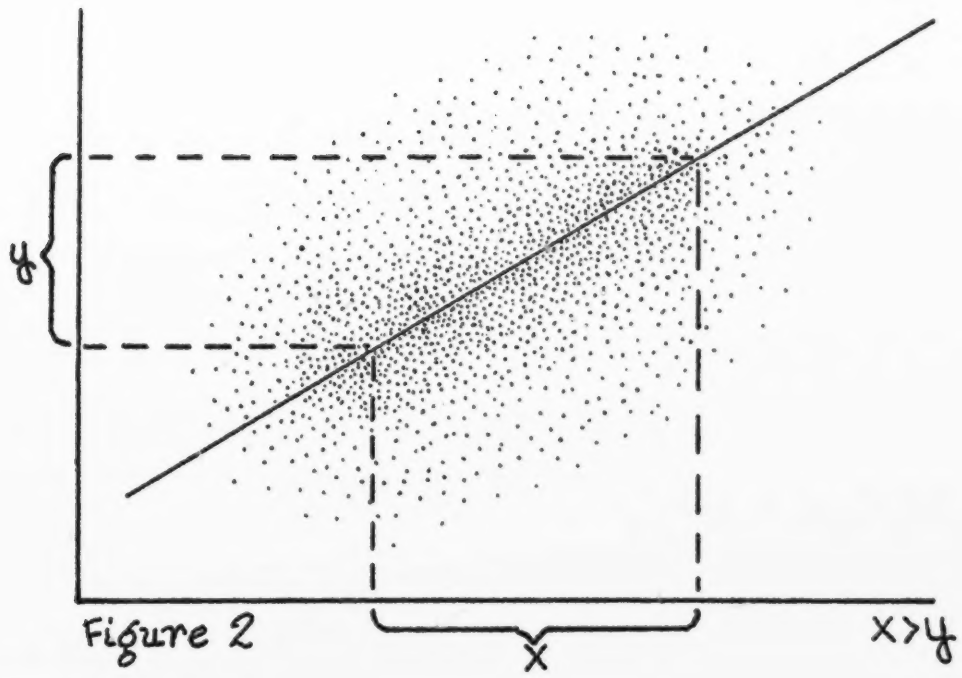

Figure 2

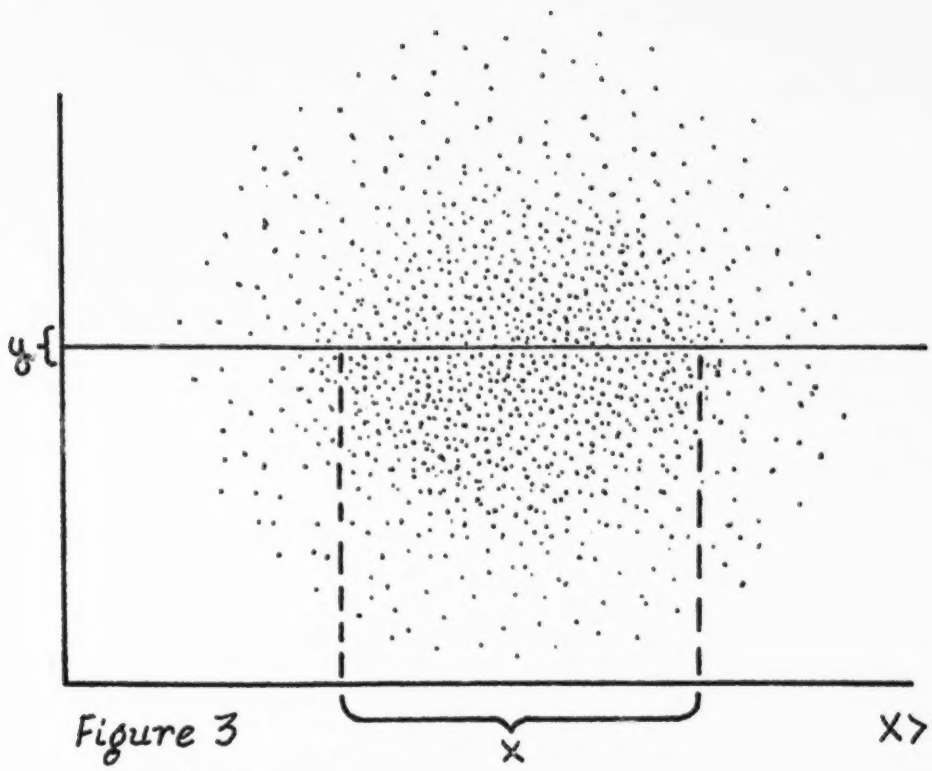

Figure 3

Sometimes the plotted data points have a distribution which clearly is non-linear (Figure 4). It is important to test for non-linearity (*curvilinear regression*).

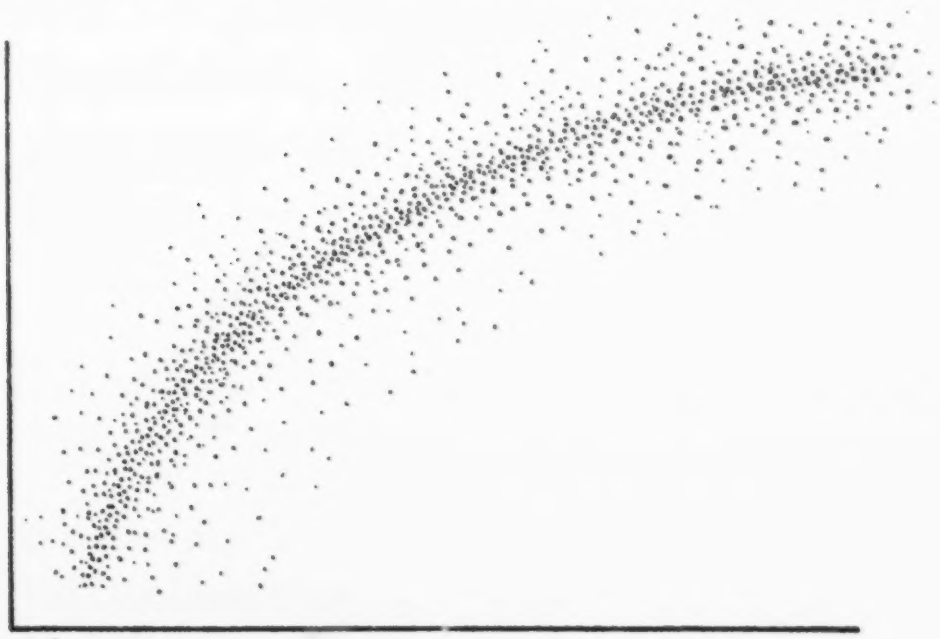

Figure 4

This is done by transgenerating the independent variables into their logarithamic functions (which linearizes parabolic and other exponential curves) and submitting them to the usual linear computations. A significant improvement in the correlation indicates *curvilinear correlation*.

Multiple correlation will be familiar to the layman if he will recall a situation where the phrase, "It depends upon several factors" would have been an accurate description. It is important to note that all factors do not have the same weight. A multiple regression equation mathematically describes the line of least squares fit after assessing the contribution of each factor. The essence of such an equation is the set of coefficients which gives appropriate weighting to each factor to be used in prediction. If the equation is calculated stepwise, one additional factor is added at each step and the cumulative multiple correlation coefficient is improved. The law of diminishing returns is in evidence and a decision regarding feasibility of continuing is required.

Appendix II

ENGLISH PLACEMENT AT EL CAMINO COLLEGE

In spring 1967, 376 students took English placement tests consisting of the Purdue Test and the Cooperative English Test. Some of these testees enrolled in their initial English course during the summer session, which was considered atypical for grading purposes. Of testees enrolled in English A or English 1A as their initial English course during fall 1967 semester and received a letter grade, only 95 had high school transcripts on file and had taken the School and College Ability Test (SCAT). Of these, 48 were placed in English A; the remaining 47 were placed in English 1A, using the Cooperative English Test total score of 157 as a cutoff score.

During the summer of 1967, of the 3,528 students who were tested via the Cooperative English Test, scores were available on the locally-prepared essay test for only 2,840 students. This number was further reduced to 1,321 by eliminating those who either had not taken the SCAT, or who did not enroll and complete an initial English course with a letter grade. Of these, 640 completed English A with a letter grade, 681 completed English 1A.

No one student, much less a population of students, can be identified as having completed the Purdue Test, the Cooperative English Test, the SCAT, and the essay test. Therefore, if either the Purdue Test or the essay test shows promise as a predictor of success in English A or 1A, the analyses performed as the focal point of this report will be inconclusive. It will remain for an acceptable size sample of entering students to take all of the tests, subsequently enroll in English A or 1A, and receive a letter grade. Then, and only then, can a multiple regression equation be generated incorporating all the appropriate variables. This equation can then predict success or lack thereof in English A or 1A, resulting in more appropriate placement in initial English courses.

The test scores were reduced to the individual section and subsection scores and were submitted to the computer along with any cumulative or total scores which may have already been available on punched cards. Non-linear transformation was performed on all independent variables for possible inclusion in the generated regression equation. Standard scores were used to eliminate anomalies created by different forms of the same test. The data were submitted to the computer in the following form:

Spring Testees

Purdue Test -- 7 sectional scores plus Total score

Coop. English Test -- 2 sectional scores plus Total score

High School Grade Point Averages -- English courses only
Overall

SCAT -- Verbal
Quantitative
Total score

Summer Testees

Coop. English Test -- 3 sectional scores (V, L, S)
Reading score (total of V, L, S)
Expression score
Total score

SCAT -- Verbal
Quantitative
Total score

Essay

The letter grade received in English A or 1A in fall 1967 was in all cases the dependent variable.

RESULTS

The computer ranked the following test scores in order of their effectiveness as predictors:

English A

Spring Testing	Cumulative Multiple Correlation Coefficient	Summer Testing	Cumulative Multiple Correlation Coefficient
High School English GPA	0.540	Essay	0.183
Purdue Test, section 7	0.634	Coop. Eng. Test, section E	0.240
Purdue Test, section 6	0.724	SCAT, Total score	0.282
Purdue Test, section 3	0.758	Coop. Eng. Test, section R	0.327
Purdue Test, section 4	0.771	Coop. Eng. Test, Total score	0.331
Purdue Test, section 5	0.792	SCAT, Quantitative section	0.345
Purdue Test, section 1	0.798		
Purdue Test, Total score	0.801		
Purdue Test, section 2	0.806		
Coop. Eng. Test, section E	0.811		
SCAT, Total score	0.816		
High School Overall GPA	0.834		
Coop. Eng. Test, Total score	0.836		
Coop. Eng. Test, section R	0.848		
SCAT, Verbal section	0.850		
SCAT, Quantitative section	0.855		

English 1A

Spring Testing	Cumulative Multiple Correlation Coefficient	Summer Testing	Cumulative Multiple Correlation Coefficient
Purdue Test, section 6	0.558	Coop. Eng. Test, section E	0.361
High School English GPA	0.638	SCAT, Verbal section	0.403
Purdue Test, section 3	0.659	Essay	0.429
Coop. Eng. Test, section R	0.685	Coop. Eng. Test, section V	0.436
SCAT, Total score	0.703	SCAT, Total score	0.437
Purdue Test, section 5	0.725	SCAT, Quantitative score	0.442
High School Overall GPA	0.744	Coop. Eng. Test, section R	0.443
SCAT, Verbal score	0.752	Coop. Eng. Test, Total score	0.445
Purdue Test, section 4	0.830		
Purdue Test, section 7	0.861		
SCAT, Quantitative section	0.880		
Coop. Eng. Test, Total score	0.883		
Coop. Eng. Test, section E	0.885		
Purdue Test, section 2	0.887		
Purdue Test, Total score	0.888		
Purdue Test, section 1	0.905		

Initial inspection of the computer-ranked predictors indicates that the Purdue Test, essay, and high school English grades show most promise as predictors of academic performance in initial English courses. The relative merits of the Purdue Test and essay cannot be assessed until a population of students has been subjected to both tests and subsequently complete their initial English course with a letter grade.

Deserving of mention is that if the proposed method is applied and high school grades are required, high school transcripts will be required to be on file for advanced placement in English. Clerical manpower will be required to evaluate these transcripts and reduce derived data to a punched card.

Correlation was improved (from 0.05 minimum to 0.19 maximum) by transgenerating the input data in order to investigate possible curvilinear correlation. Concern was expressed that selection of both a score and its transgenerated value by the computer would indicate that the data was both linear and non-linear simultaneously which is impossible. Continued consideration will reveal that the data can indeed *approach* both descriptions via least squares fit, especially since the transgenerated values may themselves approach linearity if the region of consideration is narrow relative to the absolute value of the untransgenerated variable. Selection of a variable in its remaining form after its initial selection does indicate that it is a better predictor than the remaining unselected variables.

It is noteworthy that the criterion now used for English placement at El Camino College (Cooperative English Test, total score) never placed

higher than seventh in any ranked list. Its correlation with actual academic performance in English courses ranged from 0.140 to 0.356.

RECOMMENDATIONS

It is recommended that:

1. A number of testing sections be randomly selected so as to cover the entire testing period rather uniformly and in adequate number to produce a sample of at least 100 students completing English A with a letter grade other than W, I, or NADF, and a like sample completing English 1A.

2. These testing sections be given the Cooperative English Test, the Purdue Test, *and* the essay (205 minutes of testing).

3. The results of each of these tests be entered *section by section plus total scores* on a single punched card. This card will contain, in addition to identifying data, SCAT scores and two high school grade point averages (overall and English courses).

4. At the end of the next semester, a final letter grade in English for each surviving student of the sample is added to the punched card.

5. A stepwise regression (multiple correlation) be performed on the raw data (including non-linear transgeneration). The final letter grade received in the course shall be the dependent variable. After selection of the appropriate variables considered feasible for future testing, the two appropriate regression equations generated shall be used in turn to predict degree of success in English of students tested in the corresponding test period one year later. Only the applicable test sections are given. Selected testing sessions receive all tests to begin a new cycle per Recommendation 1 (above). Prediction of lack of success in English 1A will result in tentative placement in English A with application of its corresponding regression equation. Prediction of lack of success in English A will result in placing the student in English R. These equations can be stored and used in the locally operated computer.

6. The Purdue Test be administered to the major mass of testees during the summer testing for fall 1968 English placement. The multiple regression equations developed on the spring 1967 testing population will suffer lack of representativeness and from small sample size, but appear to offer a better criterion than the single total score of the Cooperative English Test. It is recommended that these equations be used for placement with somewhat higher academic performance cutoff scores to offset the limited range-of-talent effect inherent in exercising placement *via* the prior criterion. A study should be made of the ramifications upon English 1A, A, and R enrollments of various academic performance cutoff scores.